Induction Hit Finding Update:
Low Energy Hit Finding Performance Studies

Klaudia Wawrowska
K.Wawrowska@sussex.ac.uk

University of Sussex/ Rutherford Appleton Laboratory

05/07/2022

Absolute Running Sum (absRS) recap

Developed the absRS algorithm which effectively integrates the incoming signal in real time, making it easier to identify trigger primitives online.

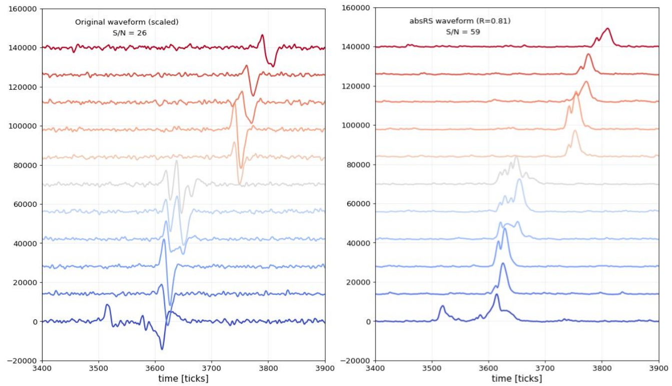
$$I_{RS}(n) = R \cdot I_{RS}(n-1) + \frac{|I_{raw}(n)|}{s}$$

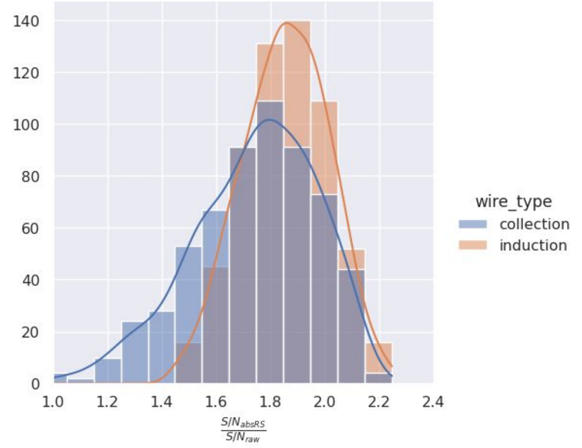
R = 0.7 (weighting factor)
s = 2 (scaling factor)

I_{RAW} = ADC value at time t
I_{RS} = running sum value at time t

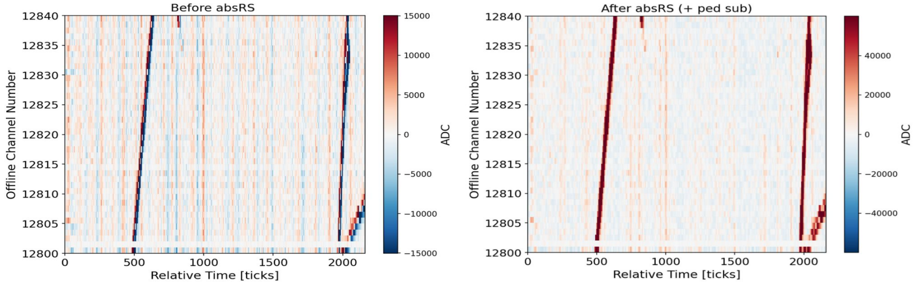
Modulated waveforms strictly unipolar.
Works for collection & induction signals.
S/N ratio increases by a factor of ~2.

Simulated data





protoDUNE data



Overview of current work

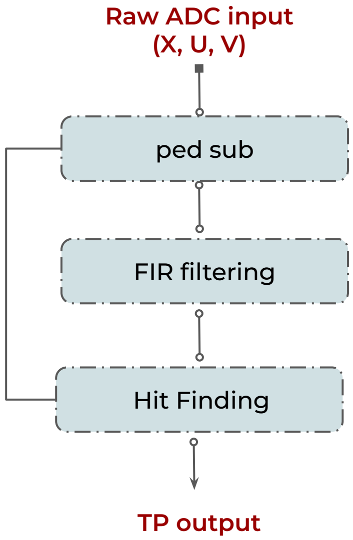
Aimed to determine whether this S/N improvement translates to low energy hit finding:

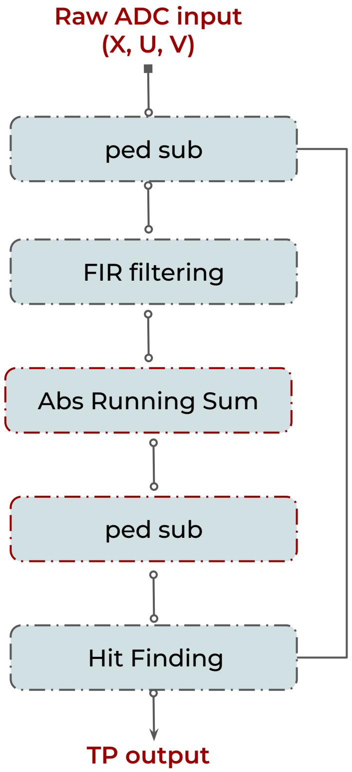
a) Generated single particles close to DUNE's energy threshold:
   i) p, e, μ samples
   ii) 500 events/particle
   iii) E_K = [5-100] MeV
b) Passed the resulting detsim waveforms through two HF chains: standard HF & absRS HF.
c) Compared the predicted hit data and the true (noiseless) hit data using a Confusion Matrix.
d) For each HF, scanned through various thresholds such that maximal sensitivity is reached before running into noise.
e) Once the threshold is optimised, directly compared the performance of the two HF algorithms to determine whether absRS is indeed more sensitive to low energy induction signals.

Hit Finding Algorithm Chains

The two types of HF chains compared in this study :

Standard HF





absRS HF *

*both extra steps based on lightweight, one-pass algorithms → shouldn't burden the firmware.

Performance metrics

Performance of different HF algorithms measured using the **confusion matrix (CM).**

CM quantifies how likely it is for a given HF algorithm to confuse noise and signal hits, and allows us to define set of useful metrics:

		True Hits	
		Yes	**No**
Predicted hits	**Yes**	True Positive (TP)	False Positive (FP)
	No	False Negative (FN)	True Negative (TN)

1: Accuracy = (TP+TN)/ (TP+FP+FN+TN)
- Ratio of correctly predicted number of hits to total number of observed hits.

2: Precision = TP/(TP+FP)
- Ratio of correctly predicted hits to total number of predicted hits.

3: Sensitivity = TP/(TP+FN)
- Ratio of correctly predicted hits to true number of hits in the event.

4: F1 score = 2*(Sensitivity * Precision)/ (Sensitivity + Precision)
- Weighted average of Precision & Sensitivity → takes false positives and false negatives into account.

Preparing data samples for analysis

Data samples for the true and predicted classes at the detsim stage:

Noiseless waveforms

Add no noise

standard HF
(thresh = 10 ADC)

True standard hits

absRS HF
(thresh = 10 ADC)

True absRS hits

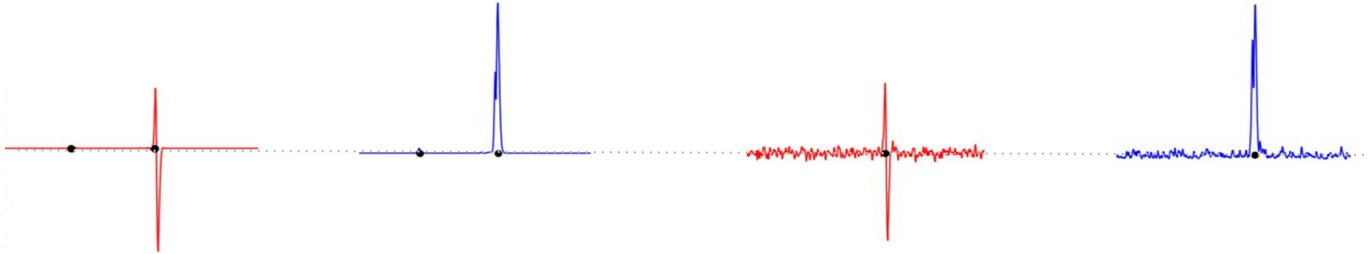
Add noise

standard HF*

Predicted standard hits

absRS HF*

Predicted absRS hits



True class

Predicted class

*first 100 ticks on each ch. ignored so pedestal can stabilize.

Hit Matching & threshold minimisation

Compared the predicted and true classes via hit matching:

True & predicted hits matched if :

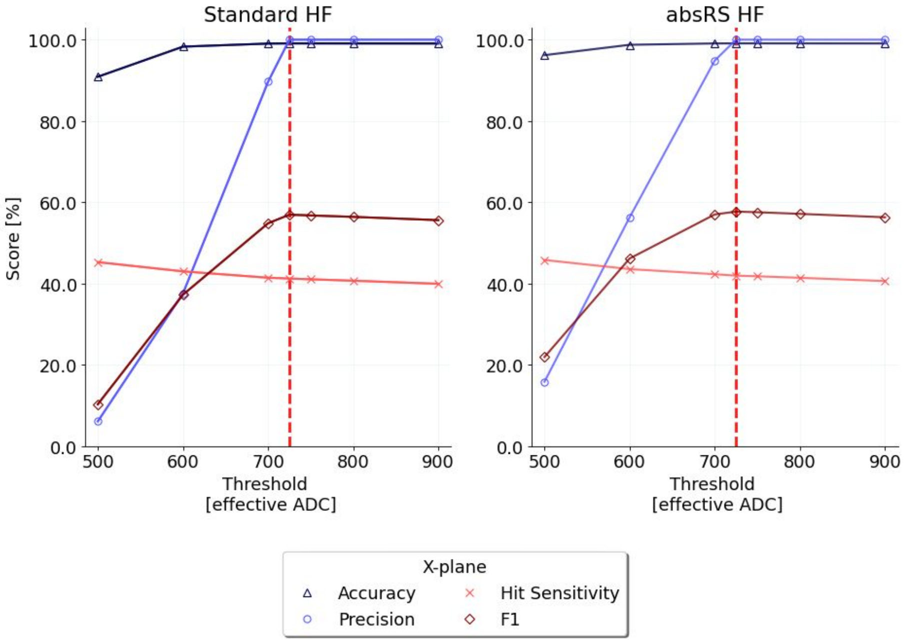
1) Both occur on the same channel
2) Hit time windows overlap

Used TP, TN, FP and FN to estimate the accuracy, precision, sensitivity & F1 for a given threshold.

Located minimum threshold achievable by doing a threshold scan:

- Low E threshold optimized when F1 is maximised
- Corresponds to best achievable sensitivity before precision begins to degrade.

Once the thresholds are optimised, I could directly compare the two HF chains for each particle type.



-- Optimised threshold for both hit finders = 725 ADC.

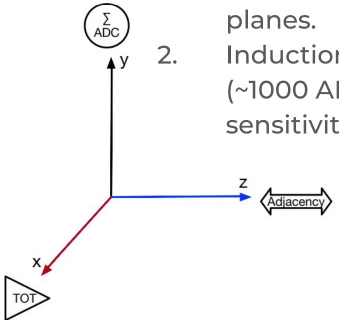
Noiseless hit spectra before and after

Main differences:

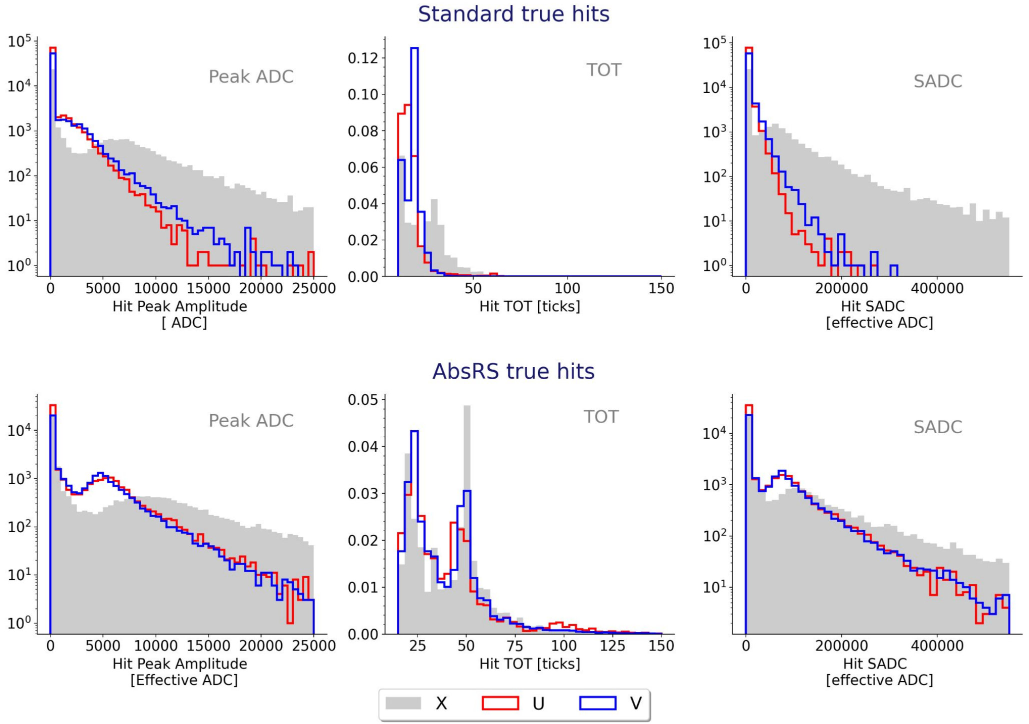
1. absRS hits are amplified
   a. right shifted peak ADC spectrum
   b. extended SADC spectrum.
2. absRS hits are elongated:
   a. ind. hits now constitute of combined -ve and +ve peaks.
   b. Overlapping ind. hits less likely to cancel out and fall below threshold.
   c. TOT slightly extended by the running sum.

Main advantages:

1. Matching X,U,V hit spectra in 3D → better directional HF performance in induction planes.
2. Induction ionization peak above noise (~1000 ADC) → expect improved low Q hit sensitivity.

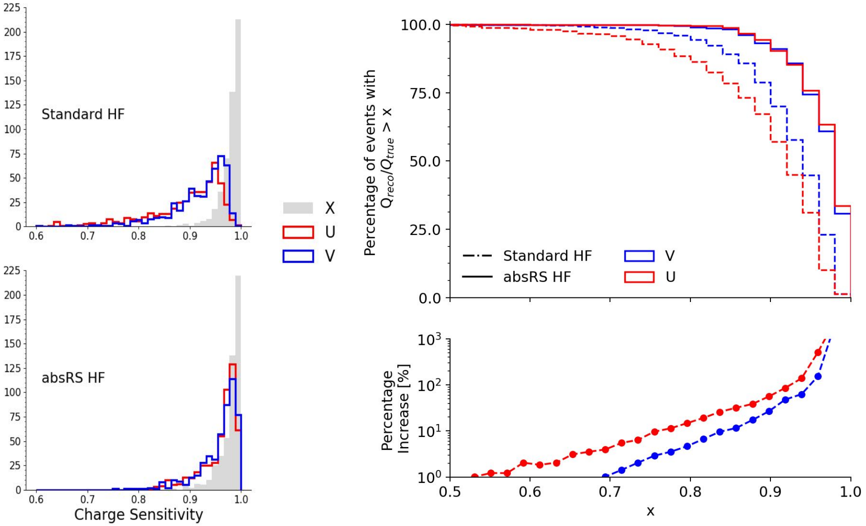


5 - 100 MeV noiseless electron hits (thresh > 10 ADC)

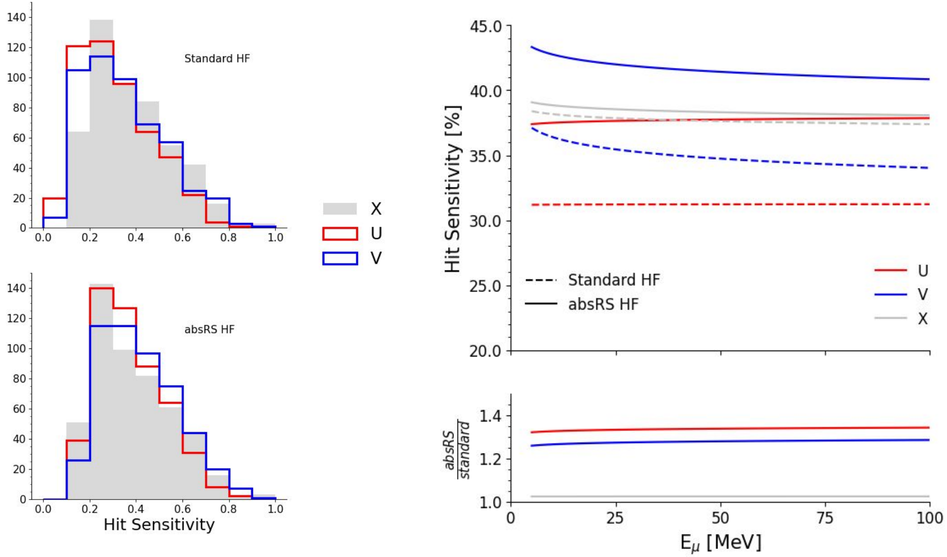


Results: 5 - 100 MeV muons

Charge Selection Efficiency



Hit Selection Efficiency



Fraction of events with $Q_{predicted}/Q_{true}$ > 0.9 (0.95)

	X	U	V
Standard HF	0.98 (0.89)	0.52 (0.18)	0.70 (0.35)
absRS HF	0.98 (0.90)	0.86 (0.67)	0.91 (0.71)
Improvement	N/A	65% (280%)	30% (100%)

Average number of hits found per event increases by 3% (X), 35% (U) & 28% (V).

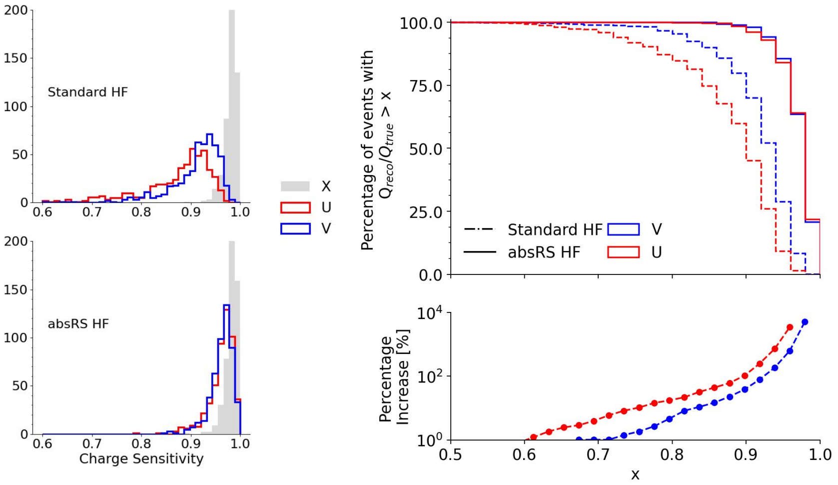
Observe much fewer events with very low hit multiplicity in induction planes → improved induction - collection agreement for hard-to-see events.

O(100)% increase in number of events with Q_{reco}/Q_{true} > 0.95.

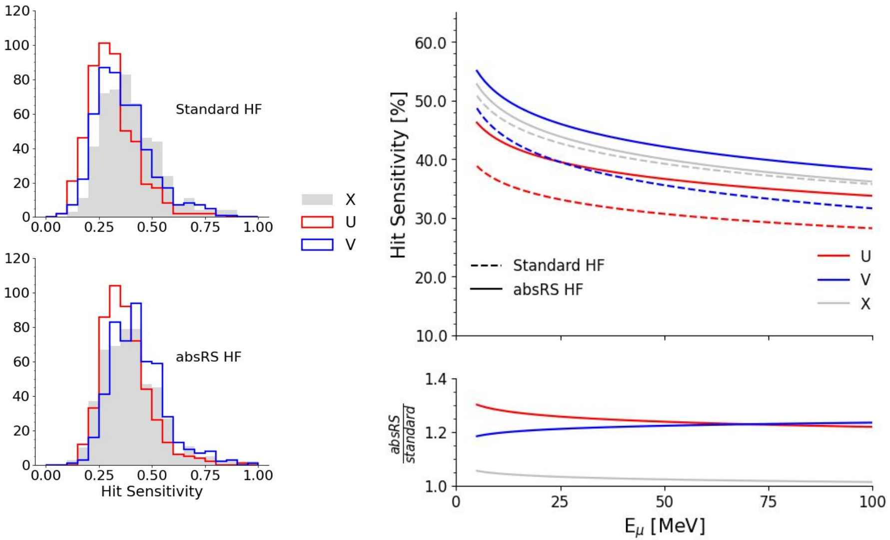
Results: 5 - 100 MeV electrons

DEEP UNDERGROUND NEUTRINO EXPERIMENT
UNIVERSITY OF SUSSEX

Charge Selection Efficiency



Hit Selection Efficiency



Fraction of events with $Q_{predicted}/Q_{true}$ > 0.9 (0.95)

	X	U	V
Standard HF	0.998 (0.97)	0.45 (0.04)	0.70 (0.18)
absRS HF	0.998 (0.98)	0.96 (0.76)	0.98 (0.78)
Improvement	N/A	113% (1,800%)	40% (326%)

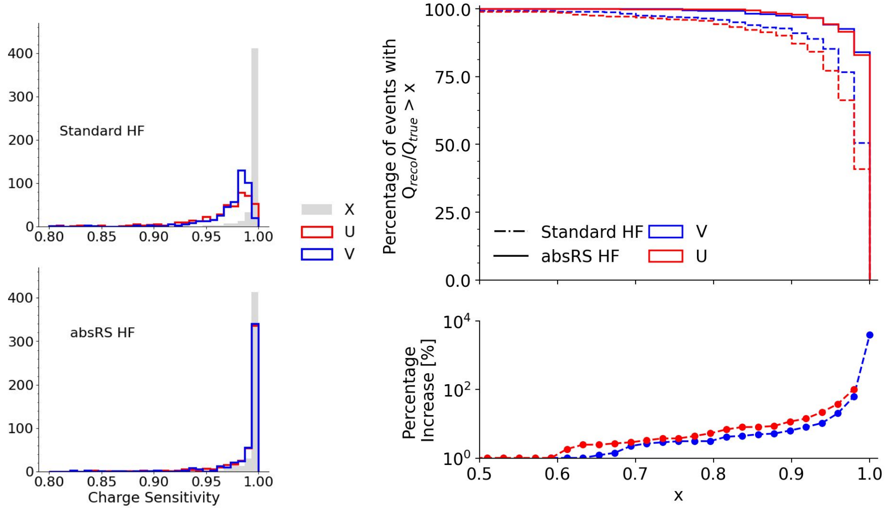
Average number of hits found per event increases by 1% (X), 24% (U) & 22% (V).

Difference in hit sensitivity between collection & induction planes effectively halved.

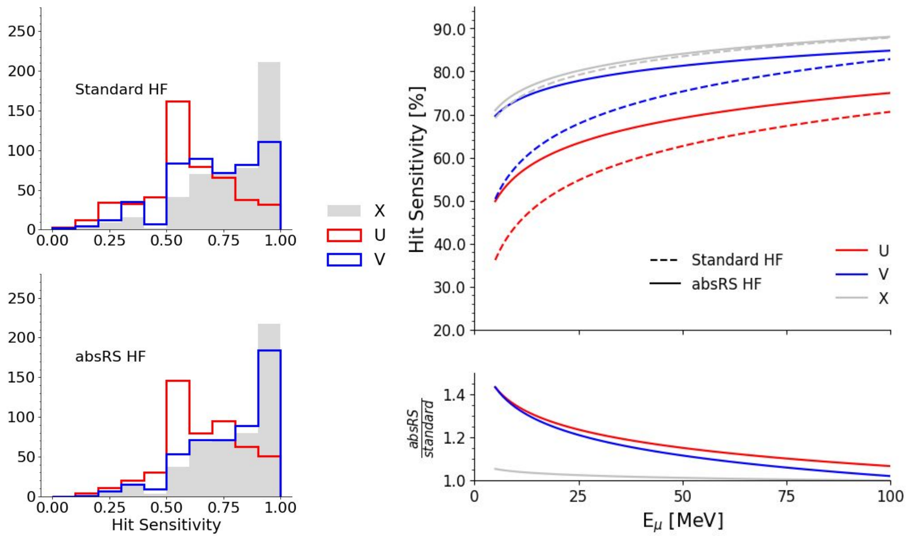
Induction Q_{reco}/Q_{true} > 0.9 for ~100% of events after absRS.

Results: 5 - 100 MeV protons

Charge Selection Efficiency



Hit Selection Efficiency



Fraction of events with $Q_{predicted}/Q_{true}$ > 0.9 (0.99)

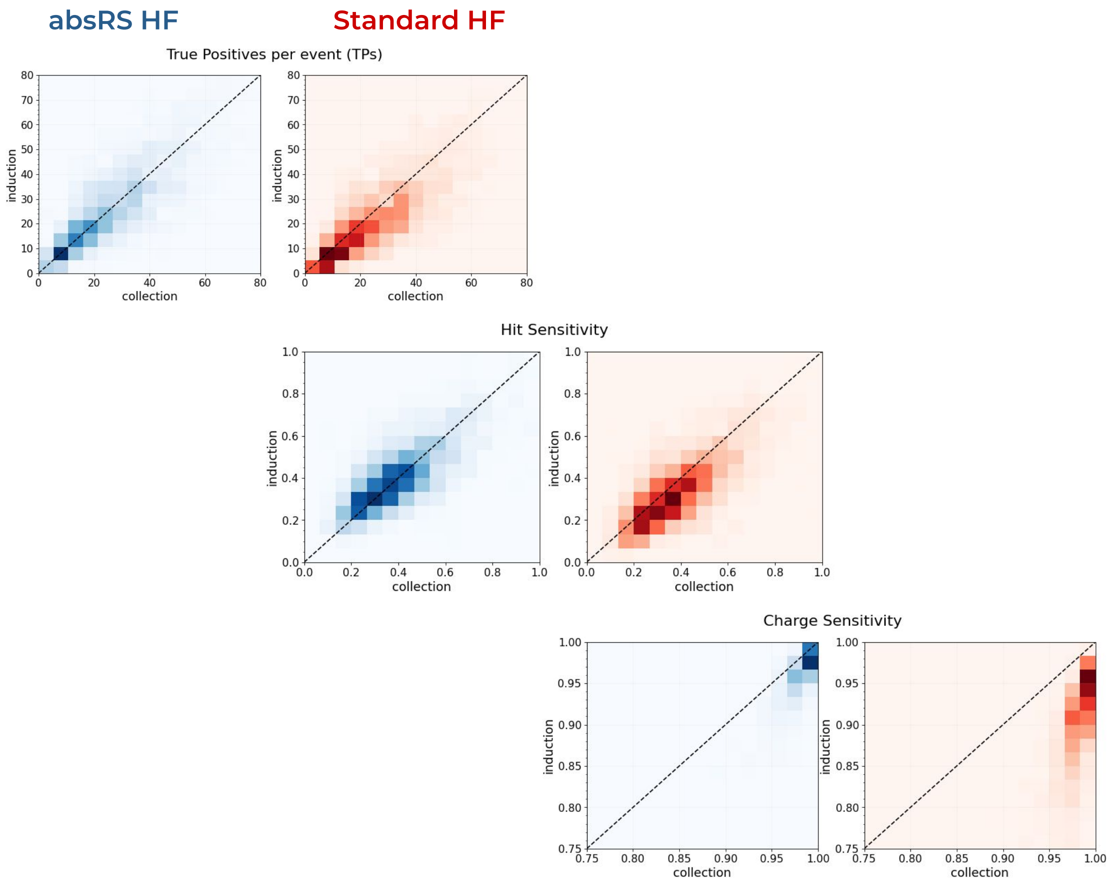
	X	U	V
Standard HF	0.98 (0.83)	0.87 (0.18)	0.91 (0.11)
absRS HF	0.98 (0.84)	0.98 (0.73)	0.97 (0.76)
Improvement	N/A	12% (304%)	7% (593%)

- Protons have relatively high Q_{dep} so charge reconstruction is relatively good in all planes (before and after).
- However, low hit multiplicity and very short track lengths (<8 cm) make hit selection challenging.
- Significant (~30%) improvement in induction hit sensitivity below 50 MeV will be benefit proton hit data selection.

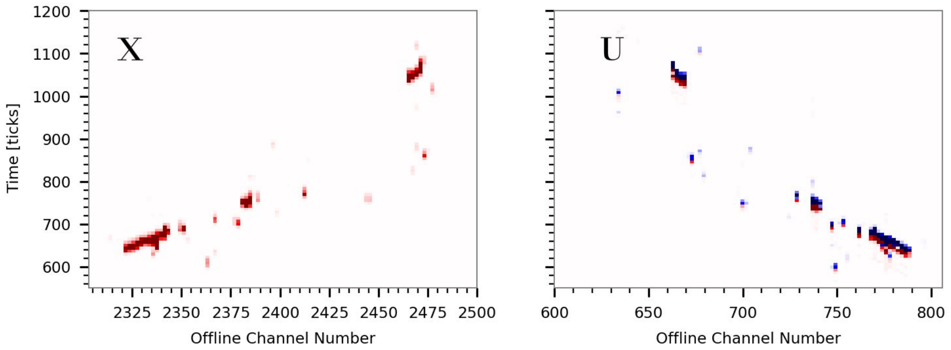
Collection-induction HF concurrence

Combined data from all events shows that after absRS we observe similarly high degree of precision and sensitivity in all three planes:

- Full 3-view redundancy at low energies.
- Matching signatures in each plane for better topological & directional reconstruction.
- Better signal significance and noise distinguishability for low Q, isolated hits (e.g. gammas, low E protons).
- Improved low energy ROI triggering.

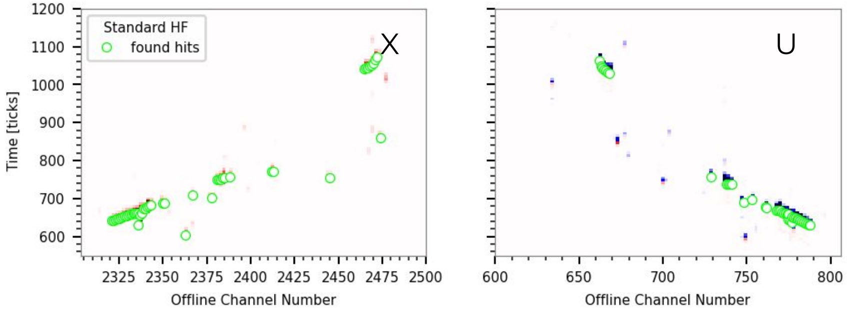


Collection-induction concurrence: event view

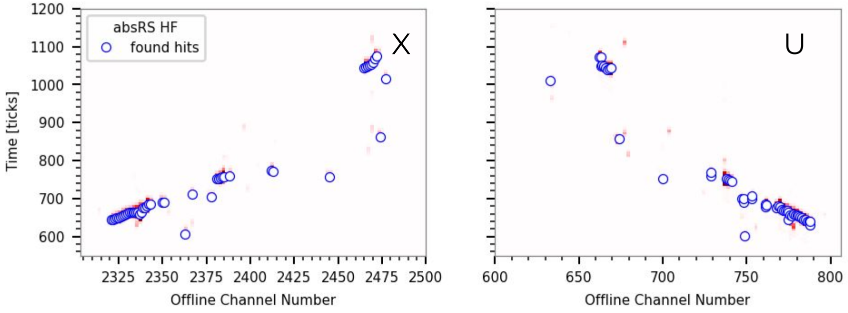


76 MeV electron track (simulated LArSoft data)

Standard HF



absRS HF



Predicted hits found using Standard HF (left) and absRS HF (right) for optimally minimised thresholds, overlaid on top of noiseless waveforms. Visibly better match in predicted shower hits for induction and collection planes after absRS.

Hit finding with ProtoDUNE-like noise

Coherent noise studies

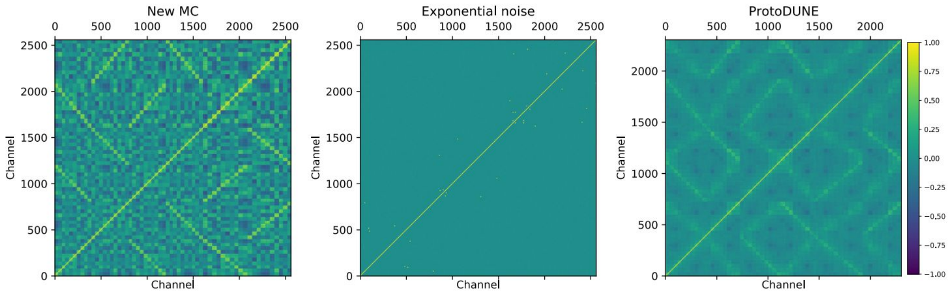
So far I've used DUNE's old "exponential noise" model.

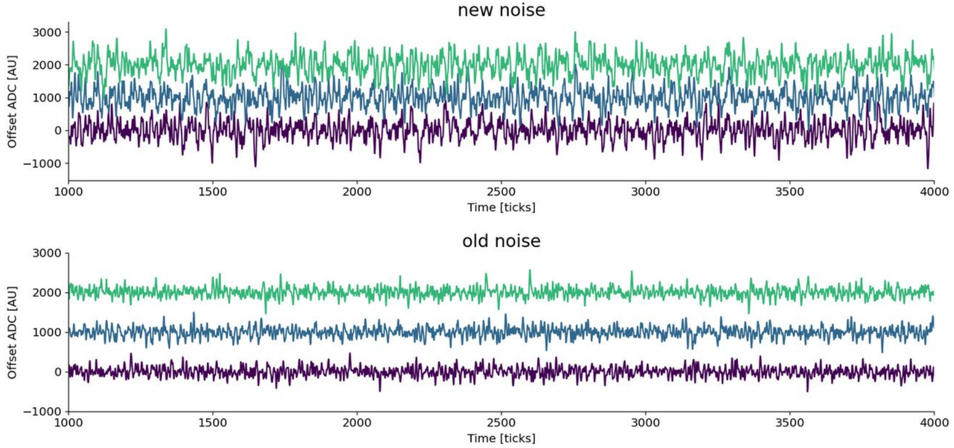
Repeated the previous study using protoDUNE-like noise ([B.Abi & P. Lasorak]) with only the electron sample (n_{eve} = 100).

New noise model consists of:

- White + coloured noise
- Digitisation noise
- Coherent noise across wires.

Aimed to establish how the RS algorithm performs in the presence of higher amplitude noise with some harmonic phase.





Threshold Scan

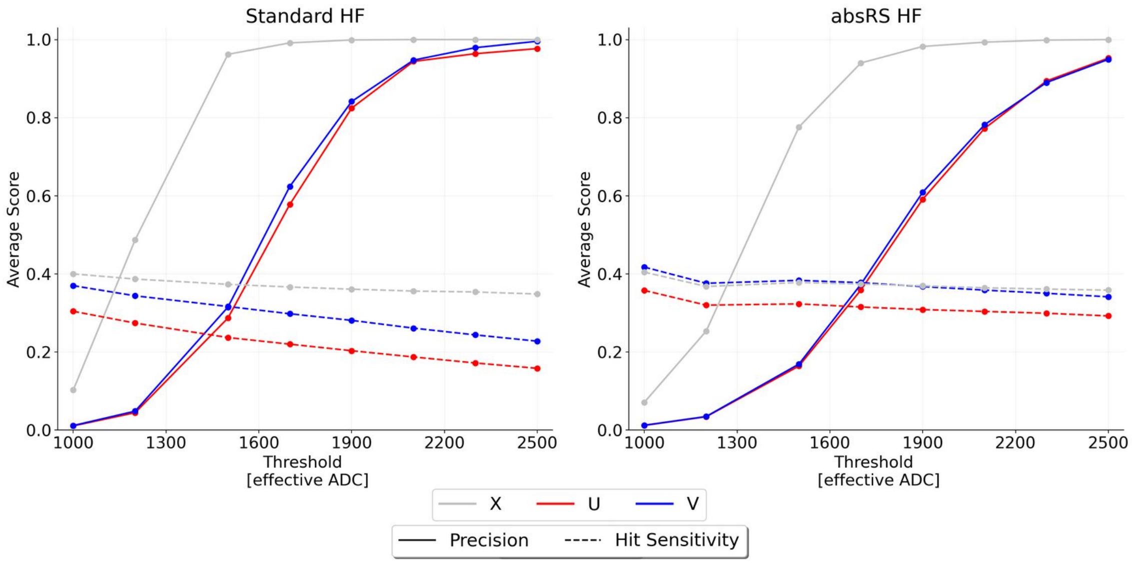
AbsRS induction hit sensitivity significantly higher & much more robust against threshold variations.

Consistently good agreement between collection and induction sensitivities.

Need slightly higher thresholds for the absRS HF to achieve high purity/precision (backup).

~Optimum performance thresholds (used for performance comparisons):

- absRS HF: 1900 (X), 2500 (U,V)
- Standard HF: 1700 (X), 2300 (U,V)



Hit Finding Performance

<u>Average Hit Sensitivity:</u>

Standard HF: 36.6% (X), 17% (U), 24% (V)

absRS HF: 36.9% (X), 29% (U), 34% (V)

→ ~40% (30%) improvement in U (V)

<u>Average Charge Sensitivity:</u>

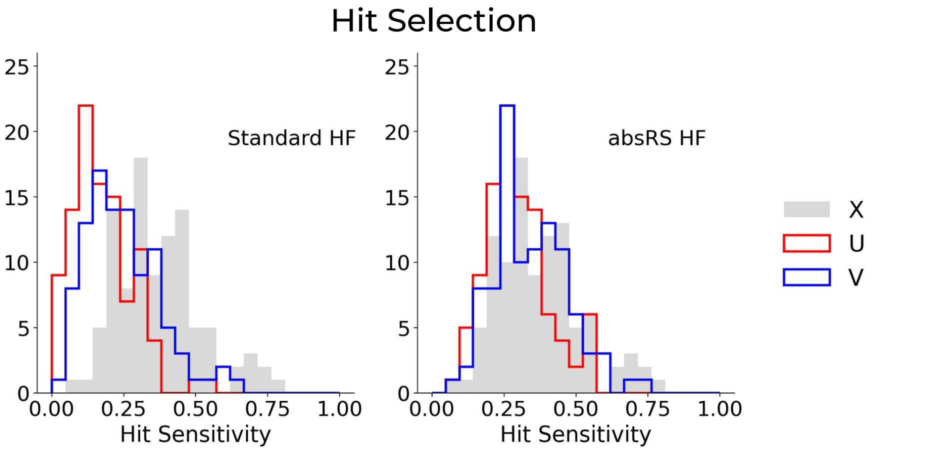
Standard HF: 96.6% (X), 61% (U), 72% (V)

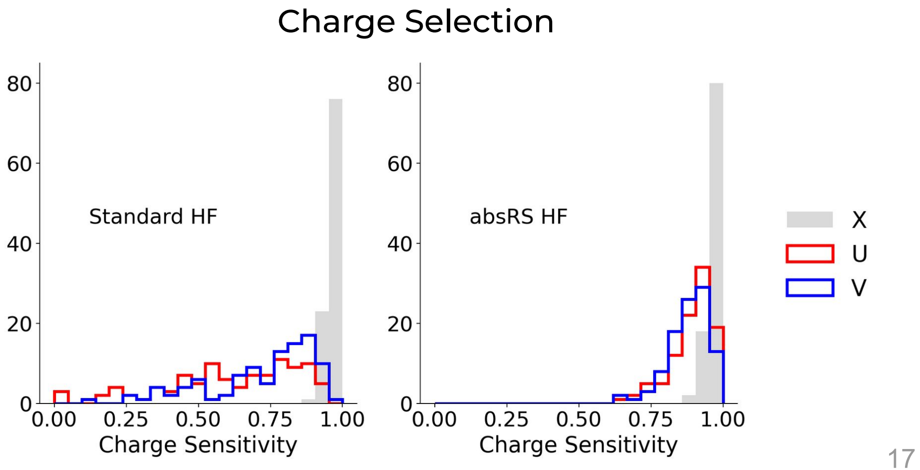
absRS HF: 96.9% (X), 89% (U), 88% (V)

→ ~30% (18%) improvement in U (V)

absRS induction HF performance only slightly worse than for the lower-noise scenario.

In the case of 'standard' induction HF, rising thresholds leads to significant degradation in performance.





Coherent noise: waveform view

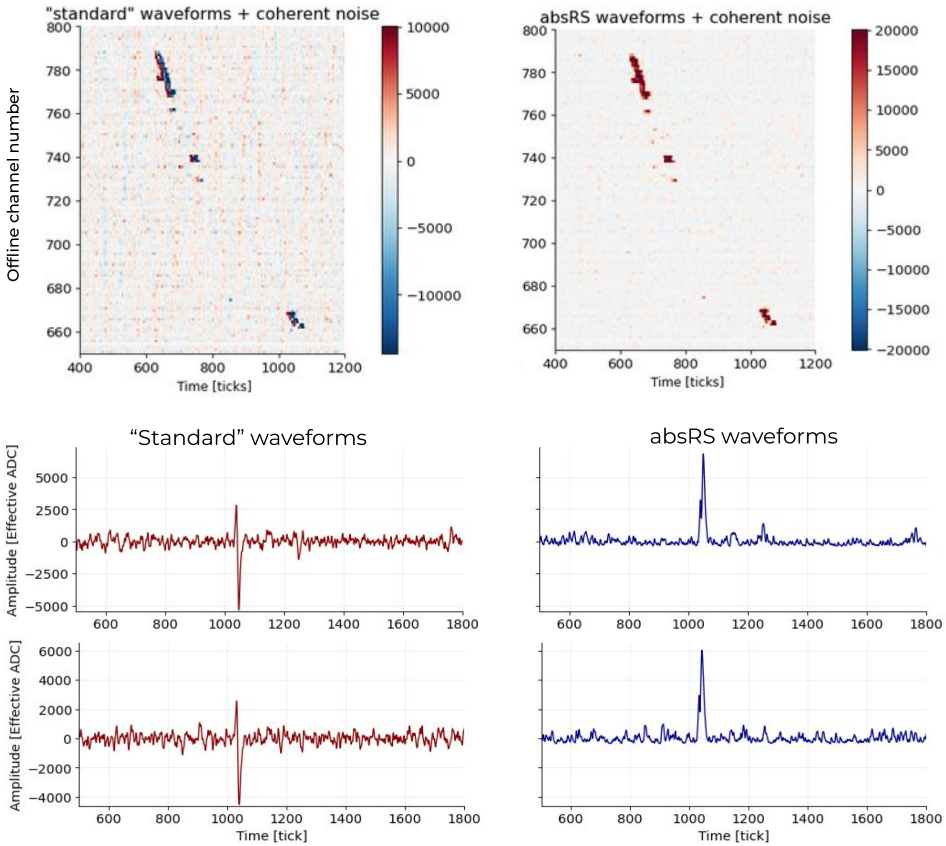
In general, running sum smoothes-out the noise.

But, any remaining higher amplitude noise spikes can get amplified leading to observable 'specks' that contribute to an increased false positive rate at intermediate thresholds.

Noise spike amplification is still very mild compared to signal amplification.

+ve peak becomes so large we can easily increase thresholds & recover the precision without sacrificing sensitivity.

→ Signal hits are still much more easily distinguishable from noise hits - especially shower hits (ch 720-760).



Conclusions

With absRS-based hit finding we can effectively reach greater sensitivities without compromising accuracy and precision.

X-, U- and V-plane hit-finding capabilities become comparable!

- Induction sensitivity becomes more robust against noise & threshold variations, making induction HF more efficient and reliable.

Having a 3-view redundancy for low Q hits can lead to improved low energy physics performance:

- Better topological 'reconstruction' online for low E events, and for awkward track orientations).
- Improved signal significance and noise distinguishability for point-like charge depositions.
- Efficient induction HF would make ROI triggering possible → allowing DUNE to lower triggering thresholds and extend it's low energy physics reach (e.g Solar & supernova neutrinos..)

Back-up slides

Hit Matching

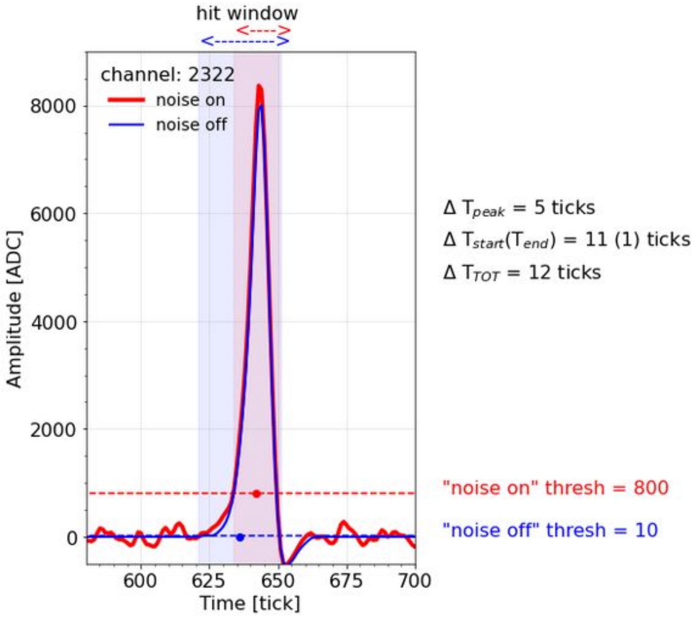
Obtain confusion matrix input parameters (TP, FN..) by comparing the predicted and true classes via hit matching:

TP++ if predicted and true hits match:

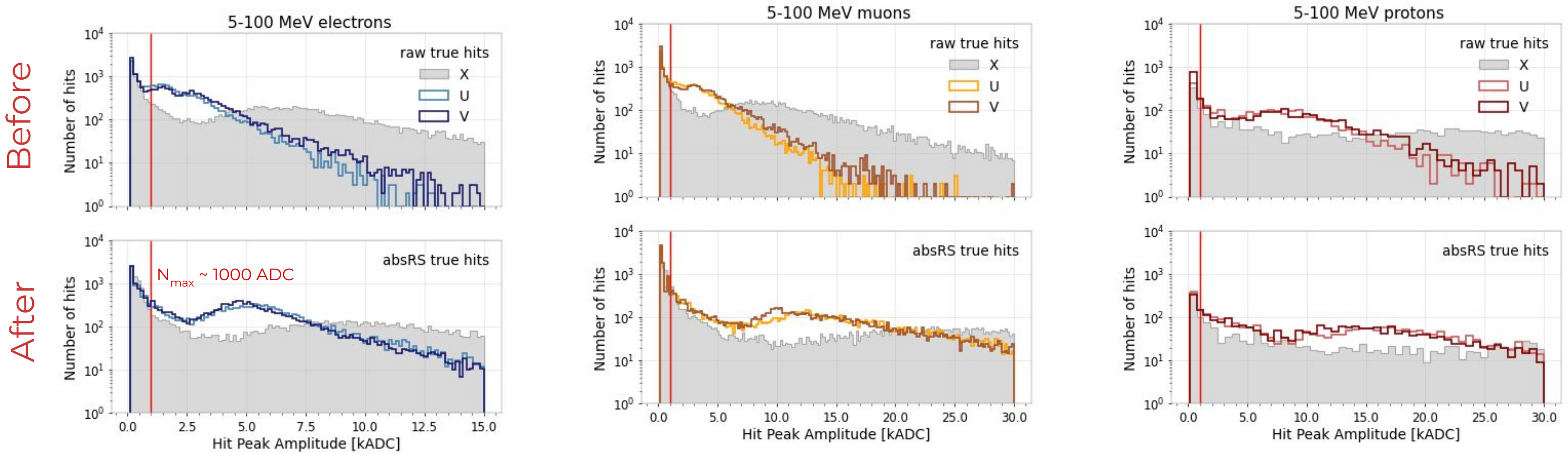
- Both hits occur on the same channel for the same event
- The hit time window of the predicted and true hit overlap

FP++ if predicted hit has no true match.

FN++ if true hit has no predicted match.



Hit/charge sensitivity trends explained



absRS amplifies signal peaks while maximum noise amplitude (N_{max}) remains unchanged (red line):

- Hit peak amplitude spectrum becomes right-shifted after algorithm is applied.
- Ionization peaks for μ & e in induction planes become clearly visible above noise → greatest improvement in sensitivity.
- Ionization peaks for protons & all three particles in X plane occur at amplitudes >> N_{max} , so absRS gives marginal improvement (the low charge peak at O(10) ADC is still invisible even after amplification).

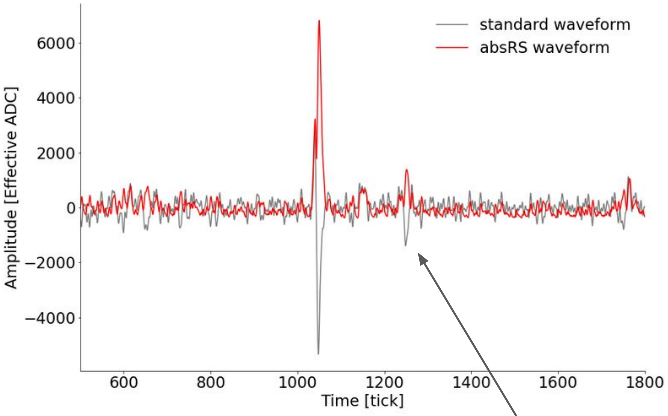
Coherent noise: Why the need for higher thresholds?

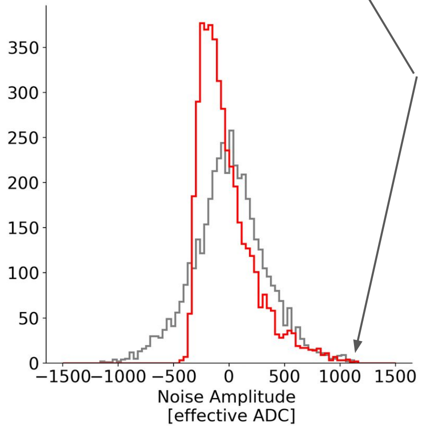
In general, absRS reduces the spread of noise due to absolute-value-based summation & scaling.

Noise probability distribution becomes right-skewed with relatively long tail towards higher ADCs.

Any low probability/ high ADC noise spikes tend to become slightly amplified → hence need for slightly higher thresholds to recover precision.

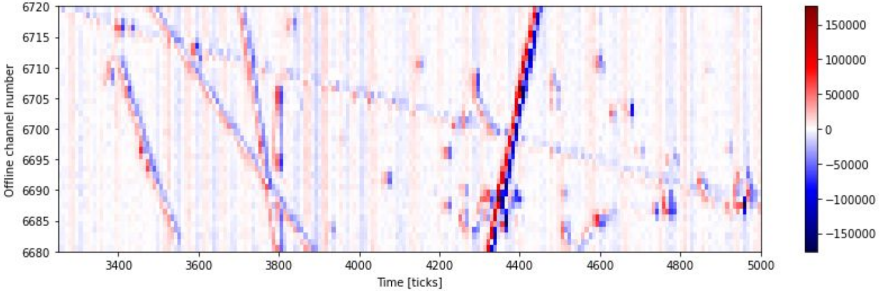
However, degree of noise amplification compared to signal amplification is tiny → hence consistently high sensitivity despite the need for higher thresholds.

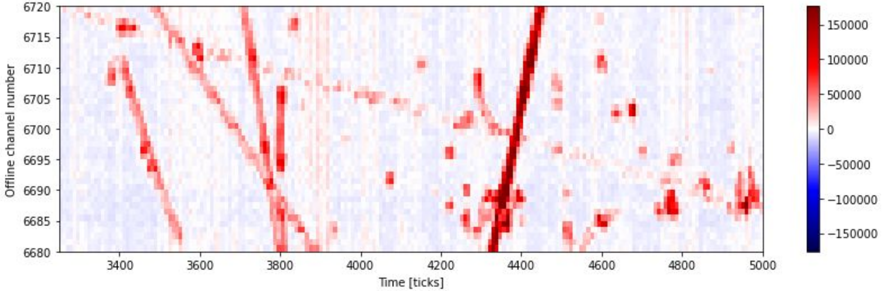




Disappearing tracks

ProtoDUNE data (run 10331, event 10)





We might be able to see disappearing tracks with absRS but they will be still very faint (the greater the amplitude - the greater the amplification).

For best results, we will definitely require some form of coherent noise removal such that thresholds can be set as low as possible.